

# SUMS OF ARITHMETIC FUNCTIONS OVER VALUES OF BINARY FORMS

R. DE LA BRETÈCHE AND T.D. BROWNING

ABSTRACT. Given a suitable arithmetic function  $h : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ , and a binary form  $F \in \mathbb{Z}[x_1, x_2]$ , we investigate the average order of  $h$  as it ranges over the values taken by  $F$ . A general upper bound is obtained for this quantity, in which the dependence upon the coefficients of  $F$  is made completely explicit.

## 1. INTRODUCTION

This paper is concerned with the average order of certain arithmetic functions, as they range over the values taken by binary forms. We shall say that a non-negative sub-multiplicative function  $h$  belongs to the class  $\mathcal{M}(A, B)$  if there exists a constant  $A$  such that  $h(p^\ell) \leq A^\ell$  for all primes  $p$  and all  $\ell \in \mathbb{N}$ , and there is a function  $B = B(\varepsilon)$  such that for any  $\varepsilon > 0$  one has  $h(n) \leq Bn^\varepsilon$  for all  $n \in \mathbb{N}$ . Let  $F \in \mathbb{Z}[x_1, x_2]$  be a non-zero binary form of degree  $d$ , such that the discriminant  $\text{disc}(F)$  is non-zero. Such a form takes the shape

$$F(x_1, x_2) = x_1^{d_1} x_2^{d_2} G(x_1, x_2), \quad (1.1)$$

for integers  $d_1, d_2 \in \{0, 1\}$ , and a non-zero binary form  $G \in \mathbb{Z}[x_1, x_2]$  of degree  $d - d_1 - d_2$ . Moreover, we may assume that  $\text{disc}(G) \neq 0$  and  $G(1, 0)G(0, 1) \neq 0$ .

Given a function  $h \in \mathcal{M}(A, B)$  and a binary form  $F$  as above, the primary goal of this paper is to bound the size of the sum

$$S(X_1, X_2; h, F) := \sum_{1 \leq n_1 \leq X_1} \sum_{1 \leq n_2 \leq X_2} h(|F(n_1, n_2)|),$$

for given  $X_1, X_2 > 0$ . For certain choices of  $h$  and  $F$  it is possible to prove an asymptotic formula for this quantity. When  $h = \tau$  is the usual divisor function, for example, Greaves [3] has shown that there is a constant  $c_F > 0$  such that

$$S(X, X; \tau, F) = c_F X^2 \log X (1 + o(1)),$$

as  $X \rightarrow \infty$ , when  $F$  is irreducible of degree  $d = 3$ . This asymptotic formula has been extended to irreducible quartic forms by Daniel [2]. When  $d \geq 5$  there are no binary forms  $F$  for which an asymptotic formula is known for  $S(X, X; \tau, F)$ . In order to illustrate the main results in this article, however, we shall derive an upper bound for  $S(X, X; \tau, F)$  of the expected order of magnitude. The primary aim of this work is to provide general upper bounds for the sum  $S(X_1, X_2; h, F)$ , in which the dependence upon the coefficients of the form  $F$  is made completely explicit. We will henceforth allow the implied constant in any estimate to depend upon the degree of the polynomial that is under consideration. Any further dependences will be indicated by an appropriate subscript.

Before introducing our main result, we first need to introduce some more notation. We shall write  $\|F\|$  for the maximum modulus of the coefficients of a binary integral form  $F$ , and we shall say that  $F$  is primitive if the greatest common divisor of its coefficients is 1. These definitions extend in an obvious way to all polynomials with integer coefficients. Given any  $m \in \mathbb{N}$ , we set

$$\varrho_F^*(m) := \frac{1}{\varphi(m)} \# \left\{ (n_1, n_2) \in (0, m]^2 : \begin{array}{l} \gcd(n_1, n_2, m) = 1 \\ F(n_1, n_2) \equiv 0 \pmod{m} \end{array} \right\}, \quad (1.2)$$

where  $\varphi$  is the usual Euler totient function. The arithmetic function  $\varrho_F^*$  is multiplicative, and has already played an important role in the work of Daniel [2]. Finally, we define

$$\psi(n) := \prod_{p|n} \left(1 + \frac{1}{p}\right), \quad (1.3)$$

and

$$\Delta_F := \psi(\text{disc}(F)). \quad (1.4)$$

We are now ready to reveal our main result.

**Theorem 1.** *Let  $h \in \mathcal{M}(A, B)$ ,  $\delta \in (0, 1)$  and let  $X_1, X_2 > 0$ . Let  $F \in \mathbb{Z}[x_1, x_2]$  be a non-zero primitive binary form of the shape (1.1). Then there exist positive constants  $c = c(A, B)$  and  $C = C(A, B, d, \delta)$  such that*

$$S(X_1, X_2; h, F) \ll_{A, B, \delta} \Delta_F^c X_1 X_2 E$$

for  $\min\{X_1, X_2\} \geq C \max\{X_1, X_2\}^{\delta d} \|F\|^\delta$ , where  $\Delta_F$  is given by (1.4) and

$$E := \prod_{d < p \leq \min\{X_1, X_2\}} \left(1 + \frac{\varrho_G^*(p)(h(p) - 1)}{p}\right) \prod_{i=1,2} \prod_{p \leq X_i} \left(1 + \frac{d_i(h(p) - 1)}{p}\right). \quad (1.5)$$

We shall see shortly that the condition  $p > d$  ensures that  $\varrho_G^*(p) < p$  in (1.5). Our initial motivation for establishing a result of the type in Theorem 1 arose in a rather different context. It turns out that Theorem 1 plays an important role in the authors' recent proof of the Manin conjecture for the growth rate of rational points of bounded height on a certain Iskovskih surface [1]. The precise result that we make use of is the following, which will be established in the subsequent section.

**Corollary 1.** *Let  $h \in \mathcal{M}(A, B)$  and let  $X_1, X_2 > 0$ . Let  $F \in \mathbb{Z}[x_1, x_2]$  be a non-zero binary form of the shape (1.1). Then we have*

$$\sum_{|n_1| \leq X_1} \sum_{|n_2| \leq X_2} h(|F(n_1, n_2)|) \ll_{A, B, \varepsilon} \|F\|^\varepsilon \left( X_1 X_2 E + \max\{X_1, X_2\}^{1+\varepsilon} \right),$$

for any  $\varepsilon > 0$ , where  $E$  is given by (1.5).

An inspection of the proof of Corollary 1 reveals that it is possible to replace the term  $X^{1+\varepsilon}$  by  $X(\log X)^{A^d-1}$ , where  $X = \max\{X_1, X_2\}$ . Moreover, it would not be difficult to extend the estimates in Theorem 1 and Corollary 1 to the more general class of arithmetic functions  $\mathcal{M}_1(A, B, \varepsilon)$  considered by Nair and Tenenbaum [7].

It is now relatively straightforward to use Theorem 1 to deduce good upper bounds for  $S(X, X; h, F)$  for various well-known multiplicative functions  $h$ . For example, on taking  $h = \tau$  in Theorem 1, and appealing to work of Daniel [2, §7] on the behaviour of the Dirichlet series  $\sum_{n=1}^{\infty} \varrho_F^*(n) n^{-s}$ , it is possible to deduce the following result, which is new for  $d \geq 5$ .

**Corollary 2.** *Let  $F \in \mathbb{Z}[x_1, x_2]$  be an irreducible binary form of degree  $d$ . Then we have  $S(X, X; \tau, F) \ll_F X^2 \log X$ .*

The primary ingredient in our work is a result due to Nair [6]. Given an arithmetic function  $h \in \mathcal{M}(A, B)$ , and a suitable polynomial  $f \in \mathbb{Z}[x]$ , Nair investigates the size of the sum

$$T(X; h, f) := \sum_{1 \leq n \leq X} h(|f(n)|),$$

for given  $X > 0$ . Nair's work has since been generalised in several directions by Nair and Tenenbaum [7]. In order to present the version of Nair's result that we shall employ, we first need to introduce some more notation. Given any polynomial  $f \in \mathbb{Z}[x]$  and any  $m \in \mathbb{N}$ , let

$$\varrho_f(m) := \#\{n \pmod{m} : f(n) \equiv 0 \pmod{m}\}.$$

It is well-known that  $\varrho_f$  is a multiplicative function. On recalling the definition (1.2) of  $\varrho_G^*(p)$ , for any binary form  $G \in \mathbb{Z}[x_1, x_2]$  and any prime  $p$ , we may therefore record the equalities

$$\varrho_G^*(p) = \begin{cases} \varrho_{G(x,1)}(p) & \text{if } p \nmid G(1,0), \\ \varrho_{G(x,1)}(p) + 1 & \text{if } p \mid G(1,0). \end{cases} \quad (1.6)$$

One may clearly swap the roles of the first and second variables in this expression. It follows from these equalities that  $\varrho_G^*(p) < p$  for any prime  $p > \deg G$ , as claimed above.

Given a positive integer  $d$  and a prime number  $p$ , we shall denote by  $\mathcal{F}_p(d)$  the class of polynomials  $f \in \mathbb{Z}[x]$  of degree  $d$ , which have no repeated roots and do not have  $p$  as a fixed prime divisor. Note that a polynomial has no repeated roots if and only if its discriminant is non-zero. Moreover, recall that a polynomial  $f \in \mathbb{Z}[x]$  is said to have fixed prime divisor  $p$  if  $p \mid f(n)$  for all  $n \in \mathbb{Z}$ . It will be convenient to abbreviate “fixed prime divisor” to “fpd” throughout this paper. When  $f$  has degree  $d$  and is primitive, then any fpd  $p$  of  $f$  satisfies  $p \leq d$ . Indeed, there are at most  $d$  roots of  $f$  modulo  $p$ . We shall write

$$\mathcal{F}(d) := \bigcap_p \mathcal{F}_p(d).$$

We are now ready to reveal the version of Nair's result that we shall employ.

**Theorem 2.** *Let  $h \in \mathcal{M}(A, B)$ , let  $f \in \mathcal{F}(d)$  and let  $\delta \in (0, 1)$ . Then there exists a constant  $C = C(A, B, d, \delta)$  such that*

$$T(X; h, f) \ll_{A, B, \delta} X \prod_{p \leq X} \left(1 - \frac{\varrho_f(p)}{p}\right) \sum_{1 \leq m \leq X} \frac{\varrho_f(m) h(m)}{m},$$

for  $X \geq C \|f\|^\delta$ .

A few remarks are in order here. First and foremost this is not quite the main result in [6, §4]. In its present form, Theorem 2 essentially amounts to a special case of a very general result due to Nair and Tenenbaum [7, Eqn. (2)]. Following our convention introduced above, the implied constant in this estimate is completely independent of the coefficients of  $f$ , depending only upon the choices of  $A, B, \delta$  and  $d$ . This uniformity will prove crucial in our deduction of Theorem 1. Theorem 2 is

in fact already implicit in the original work of Nair [6], and is a major step on the way towards his upper bound

$$T(X; h, f) \ll_{A, B, \delta, \text{disc}(f)} X \prod_{p \leq X} \left(1 - \frac{\varrho_f(p)}{p}\right) \exp\left(\sum_{p \leq X} \frac{h(p)\varrho_f(p)}{p}\right), \quad (1.7)$$

for  $X \geq C\|f\|^\delta$ . As indicated, there is now an implicit dependence upon the discriminant of the polynomial  $f$ . This arises in passing from the term  $\sum_m \frac{h(m)\varrho_f(m)}{m}$  to the term  $\exp\left(\sum_p \frac{h(p)\varrho_f(p)}{p}\right)$ .

We take this opportunity to correct an apparent oversight in recent work of Heath-Brown [4]. Here, a special case of Nair's result is used [4, Lemma 4.1], in which the dependence of the implied constant upon the polynomial's discriminant does not seem to have been accurately recorded. This leads to problems in the proof of [4, Lemma 4.2], and in particular the estimation of the sum  $S_0(m)$ , since the relevant discriminant will now vary with the choice  $m$ . Similar remarks apply to the estimation of  $S(d, d')$  in [4, Lemma 6.1]. The proof of these two estimates can now be easily repaired: the first by appealing to Theorem 2 instead of (1.7), and the second via a straightforward application of Theorem 1.

**Acknowledgement.** Part of this work was undertaken while the second author was visiting the Université de Paris-Sud, the hospitality and financial support of which is gratefully acknowledged.

## 2. PRELIMINARIES

We begin this section by establishing Corollary 1. Now it is trivial to see that  $\Delta_F \ll_\varepsilon \|F\|^\varepsilon$ , since the discriminant of a form can always be bounded in terms of the maximum modulus of its coefficients, and we have

$$\psi(n) \leq 2^{\omega(n)} \ll_\varepsilon n^\varepsilon,$$

for any  $\varepsilon > 0$ . Moreover, it will suffice to establish the result under the assumption that  $F$  is primitive. Indeed, if  $k$  is a common factor of the coefficients of  $F$ , then it may be extracted and absorbed into the factor  $\|F\|^\varepsilon$ , since  $h(ab) \ll_{B, \varepsilon} a^\varepsilon h(b)$  for  $h \in \mathcal{M}(A, B)$ . Let us take  $\delta = \varepsilon$  in the statement of Theorem 1. Suppose first that  $\min\{X_1, X_2\} \leq C \max\{X_1, X_2\}^{d\varepsilon} \|F\|^\varepsilon$ . Then since  $E \ll_\varepsilon (X_1 X_2)^\varepsilon$  in (1.5), we easily deduce that

$$S(X_1, X_2; h, F) \ll_{A, B, \varepsilon} \|F\|^\varepsilon \max\{X_1, X_2\}^{1+\varepsilon}.$$

This is satisfactory for Corollary 1. In the alternative case, Theorem 1 gives a satisfactory contribution from those  $\mathbf{n}$  for which  $n_1 n_2 \neq 0$ . The contribution from  $n_1 = 0$  is

$$\leq \sum_{|n_2| \leq X_2} h(|F(0, n_2)|) \leq h(|F(0, 1)|) \sum_{|n_2| \leq X_2} h(n_2^d) \ll_{B, \varepsilon} \|F\|^\varepsilon X_2^{1+\varepsilon},$$

since  $h \in \mathcal{M}(A, B)$ , which is also satisfactory. On arguing similarly for the contribution from  $n_2 = 0$ , we therefore complete the proof of Corollary 1.

We now collect together the preliminary facts that we shall need in our proof of Theorem 1. Let  $F \in \mathbb{Z}[\mathbf{x}]$  be a non-zero binary form of degree  $d$ . Here, as throughout our work, any boldface lowercase letter  $\mathbf{x}$  will mean an ordered pair  $(x_1, x_2)$ . If

$[\alpha_1, \beta_1], \dots, [\alpha_d, \beta_d] \in \mathbb{P}^1(\mathbb{C})$  are the  $d$  roots of  $F$  in  $\mathbb{C}$ , then the discriminant of  $F$  is defined to be

$$\text{disc}(F) := \prod_{1 \leq i < j \leq d} (\alpha_i \beta_j - \alpha_j \beta_i)^2.$$

It will be convenient to record the following well-known result.

**Lemma 1.** *Let  $\mathbf{M} \in GL_2(\mathbb{Z})$ . Then we have*

$$\text{disc}(F(\mathbf{M}\mathbf{x})) = \det(\mathbf{M})^{d(d-1)} \text{disc}(F).$$

We shall also require good upper bounds for the quantity  $\varrho_f(p^\ell)$ , for any primitive polynomial  $f \in \mathbb{Z}[x]$  and any prime power  $p^\ell$ . The following result may be found in unpublished work of Stefan Daniel, the proof of which we provide here for the sake of completeness.

**Lemma 2.** *Let  $d \in \mathbb{N}$ , let  $p$  be a prime, and let  $f \in \mathbb{Z}[x]$  be a polynomial of degree  $d$  such that  $p$  does not divide all of the coefficients of  $f$ . Then we have*

$$\varrho_f(p^\ell) \leq \min \{ dp^{\ell-1}, 2d^3 p^{(1-1/d)\ell} \},$$

for any  $\ell \in \mathbb{N}$ .

*Proof.* The upper bound  $\varrho_f(p^\ell) \leq dp^{\ell-1}$  is trivial. The second inequality is easy when  $d = 1$ , or when  $p$  divides all of the coefficients of  $f$  apart from the constant term, in which case  $\varrho_f(p^\ell) = 0$ . Thus we may proceed under the assumption that  $d \geq 2$  and  $p$  does not divide all of the coefficients in the non-constant terms. We have

$$\varrho_f(p^\ell) = \frac{1}{p^\ell} \sum_{a \pmod{p^\ell}} \sum_{b \pmod{p^\ell}} e_{p^\ell}(af(b)) = \sum_{j=0}^{\ell} \frac{1}{p^j} \sum_{\substack{a \pmod{p^j} \\ p \nmid a}} \sum_{b \pmod{p^j}} e_{p^j}(af(b)),$$

where  $e_q(z) = e^{2\pi iz/q}$ , as usual. But then the proof of [8, Theorem 7.1] implies that each inner sum is bounded by  $d^3 p^{(1-1/d)j}$  in modulus, when  $j \geq 1$ . Hence

$$\varrho_f(p^\ell) \leq 1 + d^3(1 - p^{-1}) \sum_{j=1}^{\ell} p^{(1-1/d)j} \leq d^3 p^{(1-1/d)\ell} \frac{1 - p^{-1}}{1 - p^{1/d-1}}.$$

The result then follows, since  $d \geq 2$  by assumption.  $\square$

The remainder of this section concerns the class of primitive polynomials  $f \in \mathbb{Z}[x]$  which have a *fpd*. The following result is self-evident.

**Lemma 3.** *Let  $p$  be a prime number and let  $f \in \mathbb{Z}[x]$  be a primitive polynomial which has  $p$  as a *fpd*. Then there exists an integer  $e \geq 0$  and polynomials  $q, r \in \mathbb{Z}[x]$ , such that*

$$f(x) = (x^p - x)q(x) + pr(x), \tag{2.1}$$

where  $q(x) = \sum_{j=0}^e a_j x^j$  for integers  $0 \leq a_j < p$  such that  $a_e \neq 0$ .

Our next result examines the effect of making the change of variables  $x \mapsto px + k$ , for integers  $0 \leq k < p$ .

**Lemma 4.** *Let  $p$  be a prime number and let  $f \in \mathbb{Z}[x]$  be a primitive polynomial of the shape (2.1). Then for each  $0 \leq k < p$ , there exists  $\nu_k \in \mathbb{Z}$  such that:*

$$(1) \quad 0 \leq \nu_k \leq e.$$

- (2)  $f_k(x) = p^{-\nu_k-1}f(px+k) \in \mathbb{Z}[x]$  is a primitive polynomial.  
(3) Suppose that  $f_k$  has  $p$  as a fpd, and is written in the form (2.1) for suitable polynomials  $q_k, r_k$ . Then  $e \geq p-1$  and  $\deg(q_k) \leq e-p+1$ .

*Proof.* Without loss of generality we may assume that  $k=0$ . Consider the identity

$$\frac{f(px)}{p} = x(p^{p-1}x^{p-1} - 1)q(px) + r(px),$$

and let  $b_j$  be the  $j$ -th coefficient of  $r(x)$ . It is not hard to see that the coefficient of  $x^{j+1}$  in  $f(px)/p$  is equal to

$$(a_{j-p+1} - a_j)p^j + b_{j+1}p^{j+1}, \quad (2.2)$$

where we have introduced the convention that  $a_j = 0$  for each negative index  $j$ . Let  $\nu_0$  be the  $p$ -adic order of the greatest common divisor of the coefficients of the polynomial  $f(px)/p$ , and write

$$f_0(x) = \frac{f(px)}{p^{\nu_0+1}}.$$

It is clear that  $f_0$  is a primitive polynomial with integer coefficients. Moreover, if  $e_0$  denotes the smallest index  $j$  for which  $a_j \neq 0$  in  $q(x)$ , then it is not hard to deduce from (2.2) that  $\nu_0 \leq e_0$ . In particular we have  $0 \leq \nu_0 \leq e$ . This is enough to establish the first two parts of the lemma.

It remains to consider the possibility that  $f_0$  has  $p$  as a fpd. Suppose first that  $\nu_0 < e_0$ . Then  $f_0(x) \equiv g_0(x) \pmod{p}$ , with

$$g_0(x) = \sum_{\ell=0}^{\nu_0} b_\ell p^{\ell-\nu_0} x^\ell.$$

If  $g_0$  has  $p$  as a fpd, then one may write it in the form (2.1) for suitable  $q_0, r_0 \in \mathbb{Z}[x]$ . But then

$$0 \leq \deg(q_0) \leq \nu_0 - p < e_0 - p \leq e - p,$$

which is satisfactory for the final part of the lemma. Suppose now that  $\nu_0 = e_0$ . Then  $f_0(x) \equiv g_0(x) \pmod{p}$ , with

$$g_0(x) = -a_{e_0}x^{e_0+1} + \sum_{\ell=0}^{e_0} b_\ell p^{\ell-\nu_0} x^\ell.$$

Arguing as above, if  $g_0$  has  $p$  as a fpd, then one may write it in the form (2.1) for suitable  $q_0, r_0 \in \mathbb{Z}[x]$  such that

$$0 \leq \deg(q_0) = e_0 + 1 - p \leq e + 1 - p.$$

This therefore completes the proof of Lemma 4.  $\square$

Our final result combines Lemmas 3 and 4 in order to show that there is always a linear change of variables that takes a polynomial with fpd  $p$  into a polynomial which doesn't have  $p$  as a fpd.

**Lemma 5.** *Suppose that  $f \in \mathbb{Z}[x]$  is a primitive polynomial that takes the shape (2.1) and has non-zero discriminant. Then there exists a non-negative integer  $\delta \leq e$ , and positive integers  $\mu_0, \dots, \mu_\delta$  with*

$$\mu_0 + \dots + \mu_\delta \leq (e+1)^2, \quad (2.3)$$

such that the polynomial

$$g_{k_0, \dots, k_\delta}(x) = \frac{f(p^{\delta+1}x + p^\delta k_\delta + \dots + pk_1 + k_0)}{p^{\mu_0 + \dots + \mu_\delta}} \quad (2.4)$$

belongs to  $\mathcal{F}_p(d)$ , for any  $k_0, \dots, k_\delta \in \mathbb{Z} \cap [0, p)$ .

*Proof.* Our argument will be by induction on the degree  $e$  of  $q$ . We begin by noting that the degree of  $f$  is preserved under any linear transformation of the shape  $x \mapsto ax + b$ , provided that  $a \neq 0$ . Similarly, in view of Lemma 1, the discriminant will not vanish under any such transformation. Thus it suffices to show that there exists a non-negative integer  $\delta \leq e$ , and positive integers  $\mu_0, \dots, \mu_\delta$ , such that (2.3) holds and the polynomial (2.4) has integer coefficients but doesn't have  $p$  as a **fpd**.

Let  $k_0$  be any integer in the range  $0 \leq k_0 < p$ . Then it follows from Lemma 4 that there exists  $\nu_0 \in \mathbb{Z}$  such that  $0 \leq \nu_0 \leq e$  and

$$f_{k_0}(x) = p^{-\nu_0-1} f(px + k_0)$$

is a primitive polynomial with integer coefficients. If  $e < p - 1$  then the final part of this result implies that  $f_{k_0}$  does not contain  $p$  as a **fpd**, and so must belong to  $\mathcal{F}_p(d)$ . In this case, therefore, the statement of Lemma 5 holds with  $\delta = 0$ ,  $\mu_0 = \nu_0 + 1$  and  $g_{k_0} = f_{k_0}$ . This clearly takes care of the inductive base  $e = 0$ , since then  $\delta = 0$  and  $\mu_0 = 1$ . Suppose now that  $e \geq p - 1$  and  $f_{k_0}$  contains  $p$  as a **fpd**. Then  $f_{k_0}$  can be written in the form (2.1) for suitable polynomials  $q', r'$  such that  $\deg(q') = e' \leq e - p + 1$ . We may therefore apply the inductive hypothesis to conclude that there exists a non-negative integer  $\delta' \leq e'$ , and positive integers  $\mu'_0, \dots, \mu'_{\delta'}$  with

$$\mu'_0 + \dots + \mu'_{\delta'} \leq (e' + 1)^2, \quad (2.5)$$

such that the polynomial

$$\frac{f_{k_0}(p^{\delta'+1}x + p^{\delta'}k'_{\delta'} + \dots + pk'_1 + k'_0)}{p^{\mu'_0 + \dots + \mu'_{\delta'}}} = \frac{f(p^{\delta'+2}x + p^{\delta'+1}k'_{\delta'} + \dots + pk'_0 + k_0)}{p^{\mu'_0 + \dots + \mu'_{\delta'} + \nu_0 + 1}}$$

belongs to  $\mathcal{F}_p(d)$ , for any  $k_0, k'_0, \dots, k'_{\delta'} \in \mathbb{Z} \cap [0, p)$ . Let  $\delta = \delta' + 1$ , let  $k'_i = k_{i+1}$  for  $i \geq 0$ , and write

$$\mu_0 = \nu_0 + 1, \quad \mu_i = \mu'_{i-1},$$

for  $i \geq 1$ . Then it follows that  $g_{k_0, \dots, k_\delta}(x) \in \mathcal{F}_p(d)$ , in the notation of (2.4), for any  $k_0, \dots, k_\delta \in \mathbb{Z} \cap [0, p)$ . Moreover, we clearly have  $\delta \leq e' + 1 \leq e - p + 2 \leq e$ , and (2.5) gives

$$\mu_0 + \dots + \mu_\delta \leq (e - p + 2)^2 + (e + 1) \leq e^2 + e + 1 \leq (e + 1)^2.$$

Thus (2.3) also holds, which therefore completes the proof of Lemma 5.  $\square$

Suppose that  $f \in \mathbb{Z}[x]$  is a primitive polynomial that takes the shape (2.1) for some prime  $p$ , but which does not have  $q$  as a **fpd** for any prime  $q < p$ . Then for any  $a \in \mathbb{Z}$ , the linear polynomial  $p^{\delta+1}x + a$  runs over a complete set of residue classes modulo  $q$  as  $x$  does. Thus it follows from the statement of Lemma 5 that

$$g_{k_0, \dots, k_\delta} \in \bigcap_{q \leq p} \mathcal{F}_q(d),$$

for any  $k_0, \dots, k_\delta \in \mathbb{Z} \cap [0, p)$ , where the intersection is over all primes  $q \leq p$ .

## 3. PROOF OF THEOREM 1

We are now ready to proceed with the proof of Theorem 1. Suppose that  $X_2 \geq X_1 \geq 1$ , and let  $F \in \mathbb{Z}[\mathbf{x}]$  be a primitive form of the shape (1.1). Let  $d' = d - d_2$  and  $d'' = d - d_1 - d_2$ . We may therefore write

$$G(\mathbf{x}) = \sum_{j=0}^{d''} a_j x_1^{d''-j} x_2^j,$$

for  $a_j \in \mathbb{Z}$  such that  $\gcd(a_0, \dots, a_{d''}) = 1$  and  $a_0 a_{d''} \neq 0$ . We begin this section by recording the following easy result.

**Lemma 6.** *Let  $p$  be a prime. Then we have  $p \mid \text{disc}(F)$  for any  $p \mid \gcd(a_0, a_1)$ . Moreover, if  $d_2 = 1$ , then we have  $p \mid \text{disc}(F)$  for any  $p \mid a_0$ .*

*Proof.* The first fact follows on observing that the reduction of  $F$  modulo  $p$  has  $x_2^2$  as a factor if  $p \mid \gcd(a_0, a_1)$ . If  $d_2 = 1$ , then the same conclusion holds provided only that  $p \mid a_0$ . The statement of the lemma is now obvious.  $\square$

We intend to apply Theorem 2, for which we shall fix one of the variables at the outset. Let  $q_m := \gcd(a_0, a_1 m, \dots, a_{d''} m^{d''})$ , for any  $m \in \mathbb{N}$ , and define

$$f_{n_2}(x) := \frac{x^{d_1} G(x, n_2)}{q_{n_2}}.$$

Then it is clear that  $f_{n_2}$  is a primitive polynomial of degree  $d'$  with integer coefficients. Moreover, we have

$$S(X_1, X_2; h, F) \leq \sum_{1 \leq n_2 \leq X_2} h(n_2^{d_2} q_{n_2}) \left| \sum_{1 \leq n_1 \leq X_1} h(|f_{n_2}(n_1)|) \right|. \quad (3.1)$$

We now want to apply Theorem 2 to estimate the inner sum. For this we must deal with the possibility that  $f_m$  contains a  $\text{fpd}$ . Since  $f_m$  is primitive of degree  $d'$ , the only possible  $\text{fpds}$  are the primes  $p \leq d'$ .

Suppose that  $f_m$  has  $p_1 < \dots < p_r$  as  $\text{fpds}$ . We shall combine a repeated application of Lemma 5 with the observation made at the close of §2. This leads us to the conclusion that there exist non-negative integers  $\delta_1, \dots, \delta_r \leq d - 2$ , together with positive integers  $m_1, \dots, m_r \leq d^2$ , such that

$$g_\beta(x) := \frac{f_{n_2}(p_1^{\delta_1+1} \dots p_r^{\delta_r+1} x + \beta)}{p_1^{m_1} \dots p_r^{m_r}} \in \mathcal{F}(d'),$$

for any  $\beta$  modulo  $p_1^{\delta_1+1} \dots p_r^{\delta_r+1}$ . It will be convenient to write

$$\alpha := p_1^{\delta_1+1} \dots p_r^{\delta_r+1}, \quad \gamma := p_1^{m_1} \dots p_r^{m_r}.$$

Then it follows from Lemma 1 that

$$\begin{aligned} \text{disc}(g_\beta) &= \text{disc} \left( \frac{(\alpha x + \beta)^{d_1} G(\alpha x + \beta, n_2)}{\gamma q_{n_2}} \right) = \text{disc} \left( \frac{F(\alpha x + \beta, n_2)}{\gamma q_{n_2} n_2^{d_2}} \right) \\ &= \left( \frac{\alpha^d n_2^{d-2d_2}}{\gamma^2 q_{n_2}^2} \right)^{d-1} \text{disc}(F). \end{aligned} \quad (3.2)$$

Note that  $\alpha \leq d^{r(d-1)} \leq d^{d^2}$  and  $\gamma \leq d^{rd^2} \leq d^{d^3}$ . In particular there are just  $O(1)$  choices for  $\beta$  modulo  $\alpha$ , and  $h(\gamma) \ll_B 1$ .



Our investigation so far has therefore led us to the inequality

$$\sum_{1 \leq n_1 \leq X_1} h(|f_{n_2}(n_1)|) \ll_B \sum_{\alpha} \sum_{\beta \pmod{\alpha}} \sum_{1 \leq n_1 \leq X_1} h(|g_{\beta}(n_1)|), \quad (3.3)$$

in (3.1), with  $g_{\beta} \in \mathcal{F}(d')$ . It will now suffice to apply Theorem 2 to estimate the inner sum, which we henceforth denote by  $U(X_1)$ . Note that  $\|g_{\beta}\| \ll \|f_{n_2}\| \ll n_2^d \|F\| \leq X_2^d \|F\|$ . Hence it follows from Theorem 2 that for any  $\delta \in (0, 1)$  we have

$$U(X) \ll_{A,B,\delta} X \prod_{p \leq X} \left(1 - \frac{\varrho_{g_{\beta}}(p)}{p}\right) \sum_{1 \leq m \leq X} \frac{\varrho_{g_{\beta}}(m)h(m)}{m}, \quad (3.4)$$

for  $X \gg_{A,B,\delta} X_2^{\delta d} \|F\|^{\delta}$ . In estimating the right hand side of (3.4), we shall find that the result is largely independent of the choice of  $\beta$ . To simplify our exposition, therefore, it will be convenient to write  $g = g_{\beta}$  in what follows.

We begin by estimating the sum over  $m$  that appears in (3.4). On combining the sub-multiplicativity of  $h$  with the multiplicativity of  $\varrho_g$ , we see that

$$\sum_{1 \leq m \leq X} \frac{\varrho_g(m)h(m)}{m} \leq \prod_{p \leq X} \left(1 + \frac{\varrho_g(p)h(p)}{p} + \sum_{\ell \geq 2} \frac{\varrho_g(p^{\ell})h(p^{\ell})}{p^{\ell}}\right). \quad (3.5)$$

We must therefore examine the behaviour of the function  $\varrho_g(p^{\ell})$  at various prime powers  $p^{\ell}$ . This is a rather classic topic and the facts that we shall use may all be found in the book of Nagell [5], for example. Now an application of Lemma 2 reveals that

$$\varrho_g(p^{\ell}) \leq \min \{d' p^{\ell-1}, 2d'^3 p^{(1-1/d')\ell}\},$$

for any  $\ell \in \mathbb{N}$ , since  $p$  does not divide all of the coefficients of  $g$ . Moreover, it is well-known that

$$\varrho_g(p^{\ell}) \leq d',$$

if  $p \nmid \text{disc}(g)$  or if  $\ell = 1$ . In view of the fact that  $h(p^{\ell}) \leq \min\{A^{\ell}, Bp^{\ell\varepsilon}\}$ , for any  $\varepsilon > 0$ , we therefore deduce that

$$\sum_{\ell \geq 1} \frac{\varrho_g(p^{\ell})h(p^{\ell})}{p^{\ell}} \leq d' \sum_{1 \leq \ell \leq d} \frac{h(p^{\ell})p^{\ell-1}}{p^{\ell}} + 2d'^3 \sum_{\ell > d} \frac{h(p^{\ell})p^{(1-1/d')\ell}}{p^{\ell}} \ll_{A,B} \frac{1}{p}.$$

for any prime  $p \mid \text{disc}(g)$ . When  $p \nmid \text{disc}(g)$  we obtain

$$\sum_{\ell \geq 2} \frac{\varrho_g(p^{\ell})h(p^{\ell})}{p^{\ell}} \leq d' \sum_{\ell \geq 2} \frac{h(p^{\ell})}{p^{\ell}} \ll_{B,\varepsilon} p^{-2(1-\varepsilon)}.$$

Now (3.2) implies that  $\psi(\text{disc}(g)) \leq \psi(\alpha n_2 \text{disc}(F)) \ll \Delta_F \psi(n_2)$ , where  $\psi$  is given by (1.3) and  $\Delta_F$  is given by (1.4). Drawing our arguments together, therefore, we have so far shown that there is a constant  $c_1 = c_1(A, B)$  such that

$$\sum_{1 \leq m \leq X} \frac{\varrho_g(m)h(m)}{m} \ll_{A,B} \Delta_F^{c_1} \psi(n_2)^{c_1} \prod_{\substack{d < p \leq X \\ p \nmid \text{disc}(g)}} \left(1 + \frac{\varrho_g(p)h(p)}{p}\right),$$

in (3.5). Suppose now that  $p > d > d'$ . Then one has  $\varrho_g(p) = \varrho_{f_{n_2}}(p)$ . We claim that  $p \nmid q_{n_2}$  provided that  $p \nmid n_2 \text{disc}(F)$ . But this follows immediately from the fact that  $\gcd(a_0, \dots, a_{d'}) = 1$ . Hence we have

$$\varrho_g(p) = \varrho_{x^{d_1}G(x,n_2)}(p) = \varrho_{x^{d_1}G(x,1)}(p) = \varrho_{G(x,1)}(p) + d_1, \quad (3.6)$$

provided that  $p > d$  and  $p \nmid n_2 \operatorname{disc}(F)$ . We may therefore conclude that there is a constant  $c_2 = c_2(A, B) > c_1$  such that

$$\sum_{1 \leq m \leq X} \frac{\varrho_g(m)h(m)}{m} \ll_{A,B} \Delta_F^{c_2} \psi(n_2)^{c_2} \prod_{d < p \leq X} \left(1 + \frac{\varrho_{G(x,1)}(p)h(p)}{p}\right) \prod_{p \leq X} \left(1 + \frac{d_1 h(p)}{p}\right).$$

We now turn to the size of the product over  $p$  that appears in (3.4), for which we shall use the relation (3.6) for any prime  $p$  such that  $p > d$  and  $p \nmid n_2 \operatorname{disc}(F)$ . Thus there is a constant  $c_3 = c_3(A, B)$  such that

$$\begin{aligned} \prod_{p \leq X} \left(1 - \frac{\varrho_g(p)}{p}\right) &\ll \prod_{\substack{d < p \leq X \\ p \nmid n_2 \operatorname{disc}(F)}} \left(1 - \frac{\varrho_{G(x,1)}(p)}{p}\right) \prod_{\substack{p \leq X \\ p \nmid n_2 \operatorname{disc}(F)}} \left(1 - \frac{d_1}{p}\right) \\ &\ll \Delta_F^{c_3} \psi(n_2)^{c_3} \prod_{d < p \leq X} \left(1 - \frac{\varrho_{G(x,1)}(p)}{p}\right) \prod_{p \leq X} \left(1 - \frac{d_1}{p}\right). \end{aligned}$$

Let

$$E_1 := \prod_{d < p \leq X_1} \left(1 - \frac{\varrho_{G(x,1)}(p)}{p}\right) \left(1 + \frac{\varrho_{G(x,1)}(p)h(p)}{p}\right) \prod_{p \leq X_1} \left(1 - \frac{d_1}{p}\right) \left(1 + \frac{d_1 h(p)}{p}\right),$$

and set  $c_4 = c_2 + c_3$ . Then we have shown that

$$U(X_1) \ll_{A,B,\delta} \Delta_F^{c_4} \psi(n_2)^{c_4} X_1 E_1$$

in (3.4), provided that  $X_1 \gg_{A,B,\delta} X_2^{\delta d} \|F\|^\delta$ . This latter inequality holds by the assumption made in the statement of Theorem 1.

Once substituted into (3.1) and (3.3), we may conclude that

$$S(X_1, X_2; h, F) \ll_{A,B,\delta} \Delta_F^{c_4} X_1 E_1 V_{d_2}(X_2), \quad (3.7)$$

where

$$V_{d_2}(X_2) = \sum_{1 \leq n_2 \leq X_2} \psi(n_2)^{c_4} h(n_2^{d_2} q_{n_2}).$$

We shall estimate  $V_0(X_2)$  and  $V_1(X_2)$  with a further application of Theorem 2. To begin with we note that for any prime  $p$  we have

$$q_p = \begin{cases} p, & \text{if } p \mid a_0 \text{ and } p \nmid a_1, \\ 1, & \text{if } p \nmid a_0. \end{cases}$$

When  $p^2 \mid a_0$  and  $p \mid a_1$  it is clear that  $q_p$  has  $p^2$  as a factor. Lemma 6 implies that this can only happen when  $p \mid \operatorname{disc}(F)$ .

Suppose first that  $d_2 = 0$ . Then the arithmetic function  $n \mapsto \psi(n)^{c_4} h(q_n)$  satisfies the conditions of Theorem 2. Applying this result with the polynomial  $f(x) = x$ , as we clearly may, it therefore follows that there is a constant  $c_5 = c_5(A)$

such that

$$\begin{aligned}
V_0(X_2) &\ll_A X_2 \prod_{p \leq X_2} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \leq X_2 \\ p \nmid a_0}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p \leq X_2 \\ p \mid a_0}} \left(1 + \frac{h(q_p)}{p}\right) \\
&\ll_A \Delta_F^{c_5} X_2 \prod_{p \leq X_2} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \leq X_2 \\ p \nmid a_0}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p \leq X_2 \\ p \mid a_0}} \left(1 + \frac{h(p)}{p}\right) \\
&\ll_A \Delta_F^{c_5} X_2 \prod_{\substack{p \leq X_2 \\ p \mid a_0}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \leq X_2 \\ p \nmid a_0}} \left(1 + \frac{h(p)}{p}\right),
\end{aligned}$$

for  $X_2 \gg_{A,B} 1$ . Recall the identities (1.6). Then on inserting this bound into (3.7), we therefore obtain the expected bound in Theorem 1 since

$$\prod_{\substack{p \leq X_2 \\ p \mid a_0}} \left(1 - \frac{1}{p}\right) \prod_{d < p \leq X_1} \left(1 - \frac{\varrho_{G(x,1)}(p)}{p}\right) \ll \prod_{d < p \leq X_1} \left(1 - \frac{\varrho_G^*(p)}{p}\right),$$

and

$$\prod_{\substack{p \leq X_2 \\ p \mid a_0}} \left(1 + \frac{h(p)}{p}\right) \prod_{d < p \leq X_1} \left(1 + \frac{\varrho_{G(x,1)}(p)h(p)}{p}\right) \ll_\delta \prod_{d < p \leq X_1} \left(1 + \frac{\varrho_G^*(p)h(p)}{p}\right).$$

Here we have used the elementary fact that there are at most  $\delta^{-1}$  primes  $p$  such that  $p \mid a_0$  and  $p > a_0^\delta$ .

Let us now turn to the case  $d_2 = 1$ . In particular it follows from Lemma 6 that  $p \mid \text{disc}(F)$  when  $p \mid a_0$ . Now the function  $n \mapsto \psi(n)^{c_4} h(nq_n)$  again verifies the conditions of Theorem 2. Thus we deduce that there exists a constant  $c_6 = c_6(A)$  such that

$$\begin{aligned}
V_1(X_2) &\ll_A \Delta_F^{c_6} X_2 \prod_{p \leq X_2} \left(1 - \frac{d_2}{p}\right) \prod_{\substack{p \leq X_2 \\ p \nmid a_0}} \left(1 + \frac{d_2 h(p)}{p}\right) \\
&\leq \Delta_F^{c_6} X_2 \prod_{p \leq X_2} \left(1 - \frac{d_2}{p}\right) \left(1 + \frac{d_2 h(p)}{p}\right).
\end{aligned}$$

On inserting this into (3.7), we easily derive the desired upper bound. This completes the proof of Theorem 1 when  $X_2 \geq X_1 \geq 1$ . The treatment of the case in which  $X_1 \geq X_2 \geq 1$  is handled in precisely the same way, by changing the order of summation at the outset.

## REFERENCES

- [1] R. de la Bretèche and T.D. Browning, Manin's conjecture for a certain Iskovskih surface. *In preparation*, 2006.
- [2] S. Daniel, On the divisor-sum problem for binary forms. *J. Reine Angew. Math.* **507** (1999), 107–129.
- [3] G. Greaves, On the divisor-sum problem for binary cubic forms. *Acta Arith.* **17** (1970), 1–28.
- [4] D.R. Heath-Brown, Linear relations amongst sums of two squares. *Number theory and algebraic geometry*, 133–176, London Math. Soc. Lecture Note Ser. **303** CUP, 2003.
- [5] T. Nagell, *Introduction to number theory*. 2nd ed., Chelsea, 1964.
- [6] M. Nair, Multiplicative functions of polynomial values in short intervals. *Acta Arith.* **62** (1992), 257–269.

- [7] M. Nair and G. Tenenbaum, Short sums of certain arithmetic functions. *Acta Math.* **180** (1998), 119–144.
- [8] R.C. Vaughan, *The Hardy–Littlewood method*. 2nd ed., CUP, 1997.

UNIVERSITÉ PARIS-SUD, BÂTIMENT 425, 91405 ORSAY CEDEX  
*E-mail address:* `regis.de-la-breteche@math.u-psud.fr`

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW  
*E-mail address:* `t.d.browning@bristol.ac.uk`